



Australian Bureau of Statistics

6311.0.55.001 - Microdata: Employee Earnings and Jobs, Australia, 2011-12

Latest ISSUE Released at 11:30 AM (CANBERRA TIME) 18/01/2016 First Issue

Summary

Introduction

INTRODUCTION

The Australian Bureau of Statistics (ABS) has constructed experimental statistics on employee earnings and jobs in the Australian labour market for the 2011-12 financial year by integrating person (employee) level files received from the Australian Taxation Office (ATO) and business level files from the Expanded Analytical Business Longitudinal Database (EABLD). This release provides access to the sample microdata file created from the above Integrated Dataset.

For more information about the Integrated Dataset refer to the Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12 (cat. no. 6311.0).

Microdata products are the most detailed information available from a census, survey, or administrative source and generally include confidentialised unit record level information (such as responses to individual questions on a questionnaire) and data derived from responses for two or more variables. They are released with the approval of the Australian Statistician.

A 'weight' is allocated to each employee record in the sample. The weight can be considered an indication of how many employees in the relevant population are represented by each person (employee) in the sample.

ACKNOWLEDGMENT

The ABS acknowledges the support provided by the ATO in developing the employee earnings and jobs data and the partnership of the Department of Industry, Innovation and Science in developing the EABLD. The provision of data as well as ongoing assistance provided by stakeholders is essential to enable this important work to be undertaken. The enhancing of labour statistics through data integration by the ABS would not be possible without their cooperation and support.

Any discussion in this manual of data limitations or weaknesses is in the context of using the data for statistical purposes, and is not related to the ability of the data to support the ATO's core operational requirements.

AVAILABLE PRODUCTS

A test file is available from the Downloads tab to assist users in understanding the structure

of the data and to test code. This test file does not contain real data and cannot be used for analysis. The actual microdata product is available through the ABS Data Laboratory, which enables in-depth analysis using a range of statistical software packages. Further information about the ABS Data Laboratory and other general information to assist users in understanding and accessing microdata are available from the Microdata Entry Page.

Data Items

DATA ITEMS

A complete list of data items included in the Employee Earnings and Jobs microdata product is provided in an Excel workbook that can be accessed from the Downloads tab.

Data items are available at two levels, Employees and Jobs. Users intending to apply for access to the ABS Data Laboratory should ensure the data they require, and the level of detail required, are available and applicable for the intended use. The test file would be helpful for this purpose.

Age

Age of employee as at 30 June 2012 as reported on the Individual Tax Return.

Annual business turnover

The total revenue generated by a business from the provision of goods and/or services for a given accounting period (annual).

Duration of job in reference period (in weeks)

The length of time a job was held during the reference period, presented in weeks. It has been derived from start and end dates of an employee holding that job, as reported on the individual PAYG summary.

Employment size

The number of employees in a business, presented in ranges. The employment size for a business is as updated annually by the Australian Taxation Office (for non-profiled businesses) or as reported by the ABS at a point in time during the profiling process (profiled businesses).

Geography (Statistical Area Level 4)

Determined based on an employee's home address at July 2014 as reported in the Personal Income Tax Client Register, and are aligned to the Australian Statistical Geography Standard (ACGS): Volume 1 – Main Structure and Greater Capital City Statistical Areas, July 2011 (cat. no. 1270.0.55.001).

Gross payment amount per job held during the reference period

The gross amounts recorded (by businesses) on the Individual Pay As You Go summary for each job held by an employee, during the reference period.

Industry (ANZSIC)

Industry information of each employing business. It aligns with the Australian and New Zealand Standard Industrial Classification (ANZSIC), 2006 (cat. no. 1292.0). The structure of ANZSIC comprises four levels, ranging from industry division (broadest level) to industry class (finest level). In this release, industry is provided at the division level. The industry division provides a limited number of categories which give a broad overall picture of the

economy. There are 19 divisions within ANZSIC, each identified by a number that is, that is, '1' for Agriculture, forestry and fishing, '2' for Mining, '3' for Manufacturing, etc.

Job number

It is a number allocated to a Job a person held during the reference period. It can vary from 1 to 7. Where a person holds multiple jobs, the jobs are numbered in order of magnitude of their Gross payment amounts and Job value 1 will always represent the Main Job (i.e. the job with the highest reported Gross Payment Amount). For purposes of confidentiality only the first seven jobs of a person are included in the file.

Main job

Main job is defined for each employee as the job in which they received the highest gross payment amount as reported on an Individual Pay As You Go summary.

Multiple job holders

Employees who held two or more concurrent jobs during the reference period. The multiple job holder status of an employee is determined based on the date information in the Individual Pay As You Go summary. If two or more jobs were held on the same day, the employee was identified as a multiple job holder.

Number of Jobs in reference period

The number of Jobs held by an employee during the reference period - they do not have to be concurrent. The top cut off for this variable in the Employee Earnings and Jobs microdata product is 7.

Occupation in main job

Refers to the occupation sub-major group as defined by the Australian and New Zealand Standard Classification of Occupations, First Edition, Revision 1_(cat. no. 1220.0) and identified by an employee as their 'Main salary and wage occupation'. This occupation may not necessarily relate to the Main job (i.e. the job for which they received the highest gross payment amount as reported on their Individual Pay As You Go summary).

Other jobs with which a job was held concurrently

Provides a list of jobs where an employee held concurrent jobs.

Sex

Sex of employee as at 30 June 2012 as reported on the Individual Tax Return.

Total earnings from all jobs held in reference period

The gross amounts paid to employees for work done or time worked (including paid leave) during the reference period. It is the aggregate of total payments (in cash and in kind) received by each employee in all of their jobs, as reported on an Individual Tax Return.

Test File

TEST FILE

The test file does not contain real data, and cannot be used for analysis.

A test file has been created for the Employee Earnings and Jobs microdata product. The purpose of the test file is to allow researchers/analysts to become familiar with the data structure and prepare code/programs prior to applying for, or commencing, an ABS Data

Laboratory session. This aims to maximise the value of sessions by saving users' time and resources once they enter the ABS Data Laboratory environment.

The test file mimics the structure of the Employee Earnings and Jobs microdata - it has the same data items and allowed values, however, all data in the test file is false, created through a randomisation process. Proportions of values within data items in the test file will be similar to those in the real data; however, relationships between data items are not (intentionally) maintained. It is extremely unlikely that a record in the test file would match with a genuine record in the real data.

The test file is available as a free download from the Downloads tab. It can also be made available in other file formats on request, if required. For further information users should email microdata.access@abs.gov.au or telephone (02) 6252 7714.

Conditions of Use

CONDITIONS OF USE

ABS responsibilities

The Census and Statistics Act 1905 includes a legislative guarantee to respondents that their confidentiality will be protected. This is fundamental to the trust that the Australian public has in the ABS and that trust is in turn fundamental to maintaining the quality of ABS information. Without that trust, respondents may be less forthcoming or truthful in answering ABS questionnaires. For more information, see 'Avoiding inadvertent disclosure' and 'Microdata' on the web page How the ABS keeps your information confidential.

User responsibilities

Use of ABS microdata requires individual users to adhere to responsibilities that are defined under Clause 7 of the Statistics Determination 1983 under the Census and Statistics Act 1905. These responsibilities are provided in the microdata Undertaking that is signed by a Responsible Officer of each organisation prior to microdata products being released.

Conditions of sale

All ABS products and services are provided subject to the ABS Disclaimer, ABS Copyright, ABS Privacy and ABS Conditions of Sale. Any queries relating to these Conditions of Sale should be emailed to intermediary.management@abs.gov.au. The ABS Privacy Policy outlines how the ABS handles any personal information that you provide to us.

Price

Microdata access is priced according to the ABS Pricing Policy and Commonwealth Cost Recovery Guidelines. For details refer to ABS Pricing Policy on the ABS website. For microdata prices refer to the Microdata prices web page.

How to apply for access

To apply for access to the microdata, clients should read the How to Apply for Microdata web page.

Clients should familiarise themselves with the User Manual: Responsible Use of ABS CURFs before applying for access.

Australian Universities

The ABS/Universities Australia Agreement provides participating universities with access to a range of ABS products and services. This includes access to microdata.

For further information, university clients should refer to the ABS/Universities Australia Agreement web page.

FURTHER INFORMATION

The Microdata Entry page on the ABS website contains links to microdata related information to assist users in understanding and accessing microdata.

For further information about data sources data scope, linking methodology, weighting methodology, data quality see the Explanatory Notes tab. The data items list is available from the Downloads tab.

For further information about the data structure and available data items see the Data Items page. A Data items list is also available from the Downloads tab.

For further information about the Test file see the Test file page. The Test File is available from the Downloads tab.

SUPPORT

For support in the use of this product, please contact Microdata Access Strategies on 02 6252 7714 or via microdata.access@abs.gov.au.

INQUIRIES

For further information about these and related statistics, contact the National Information and Referral Service on 1300 135 070, or email client.services@abs.gov.au. The ABS Privacy Policy outlines how the ABS will handle any personal information that you provide to us.

About this Release

Employee Earnings and Jobs (EEJ) microdata is an output of the linked employer-employee data that the ABS has constructed to demonstrate the value of this linked data for statistical use.

The linked employer-employee data integrates Personal Income Tax data for 2011-12 sourced from the Australian Taxation Office with firm-level data extracted from the ABS Expanded Analytical Business Longitudinal Database. For more information about the linked data refer to the Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12 (cat. no. 6311.0) released on 11 December 2015.

The microdata file is a unit record file released via the ABS Data Laboratory constructed in a manner not likely to enable the identification of a particular person or organisation. For more

information about the ABS Data Laboratory refer to: <https://www.abs.gov.au/websitedbs/D3310114.nsf//home/Microdata%20Entry%20Page>

The microdata product is a 10% sample of the complete integrated employer-employee file, representative of the in-scope employee level records. It includes key employee variables (such as occupation for main job, earnings per job, multiple job holding) to facilitate a cross sectional analysis of employee earnings and employee jobs together with the business characteristics (such as employment size, total sales and industry) associated with the employing business for each of those jobs.

The geography on the file is at Statistical Area Level 4, the largest sub-State regions in the Main Structure of the Australian Statistical Geography Standard.

Explanatory Notes

Explanatory Notes

EXPLANATORY NOTES

1 The Employee Earnings and Jobs microdata product contains a 10% sample (at the person level) of the integrated employer-employee file.

FILE STRUCTURE

2 The structure of the microdata product is hierarchical:

1. Person (Employee)

2. Job (along with business characteristics relating to that job)

3 For persons who had a missing job record, a 'dummy' (job) record has been created to maintain the integrity of the file structure. Data items for these records have 'Not known' values if relevant, or have been given a zero value. These records are identified by the data item *Dummy Job record data flag* (DUMJOBF) having a value of 1.

4 The same applies to business data items where a job could not be linked to a business. These records are identified by the data item *No business data available flag* (NOBUSDAT) having a value of 1.

DATA SOURCES

5 Person and business level data for 2011-2012, sourced from the Personal Income Tax data and the Expanded Analytical Business Longitudinal Database respectively, were used to construct the Integrated Dataset from which the microdata product was created.

Personal Income Tax (PIT) dataset

6 The PIT dataset contains person level unit record data compiled by the Australian Taxation Office (ATO) and consists of three subsets.

- Client Register;
- Client Dataset; and

- Individual Pay As You Go (PAYG) Dataset.

7 An extract of the PIT dataset containing selected variables has been used in constructing the microdata product and the data for each individual has been linked across these three subsets using an encrypted person identifier, the Scrambled Tax File Number.

Expanded Analytical Business Longitudinal Database (EABLD)

8 The EABLD is the longitudinal business level unit record data file created by the ABS in 2015. The Integrated Dataset used an extract of the EABLD for 2011-12 containing selected variables. The linking variable between PIT dataset and EABLD extract was Australian Business Number (ABN) as issued by ATO. For further information on the data sources and the linking methodology, refer to the Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12 (cat. no. 6311.0).

SCOPE

9 This microdata product aims to represent information on all employee earnings and jobs in Australia throughout the reference period of 1 July 2011 to 30 June 2012. The scope includes:

- All persons who were an employee at any point in the reference period as recorded on either an Individual Tax Return (ITR) or an Individual PAYG summary;
- All jobs as reported in an Individual PAYG summary during the reference period; and
- All businesses which provided an Individual PAYG summary to an employee in the reference period.

COVERAGE

10 Employees who meet one of the following conditions are excluded from coverage in the microdata product.

- Employees who did not report earnings on an ITR for any of the following reasons:
 - Did not submit an ITR for any of the reasons outlined on pages 6 and 7 of the *Individual Tax Return Instructions 2012*;
 - Did not submit an ITR for any other reason; or
 - Submitted an ITR but did not report their applicable earnings.
- Employees who did not receive an Individual PAYG summary from an employer for any reason including:
 - They worked for cash in hand or other payments not recorded on an Individual PAYG summary;
 - They conducted illicit activities not recorded on Individual PAYG summaries; or
 - They did not supply their Tax File Number to their employer.

11 There were no businesses excluded on the basis of coverage.

12 For further information on scope and coverage of the Integrated Dataset, refer to the

DATA LINKING METHODOLOGY

13 The Integrated Dataset was created through a two stage process. The first stage involved linking the component files (Client Register, Client Dataset and PAYG) within the PIT dataset, and the second stage involved integrating the linked PIT dataset with the EABLD.

14 For details of the data linking process refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

DATA CLEANING

15 This Employee Earnings and Jobs microdata product is comprised, in part, of tax data supplied by the ATO to the ABS under the Taxation Administration Act 1953, which requires that the ABS only use the data for the purpose of administering the Census and Statistics Act 1905. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes, and is not related to the ability of the data to support the ATO's core operational requirements.

16 Data cleaning was undertaken on the PIT data in order to remove duplicate records, remove invalid PAYG records (jobs with less than \$1 in gross payments), and derive data items which aligned with ABS standards and classifications, where possible. Duplicate records were identified as those where all variables were identical. Demographic variables (age and sex) were checked to ensure that they were referenced to 30 June 2012. Variables such as occupation were checked to ensure that they adhered to the ABS classifications and any erroneous or invalid codes were removed.

17 For the purposes of this microdata product, minimal data cleaning was required on the EABLD extract. In creating the EABLD, transformation of source data was required to ensure that the contents adhered to the ABS standards and classifications.

18 For details of the data cleaning process refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

SAMPLE DESIGN

19 In order to mitigate risks of disclosure only a sample (10%) of the records (person level) on the Integrated Dataset have been included in the microdata product. The sample has been chosen to be representative of the key characteristics of Employees using a stratified sample design.

20 Key aspects of the sample design are:

- The sample was stratified by Statistical Area Level 4, Occupation groups at the 1 digit level and ranges of total annual employee earnings.
- Strata were constructed to have a minimum size of 100 persons.

21 A 10% simple random sample was taken from each stratum.

22 The weighting ensured there was broad representativeness at the Statistical Area Level 4 by 1 digit Occupation and Age by Sex by 1 digit Occupation levels.

23 The microdata output contains:

- 1,033,031 persons which when weighted represent 10,333,171 persons.
- 1,387,945 job records.
- 315,674 businesses (stored at the job level). This number consists of 257,045 where business information is available and 58,629 dummy businesses allocated to jobs where business information was not available. The latter can be identified by the data item 'No business data available flag' equalling 1.

WEIGHTING

Sample weights

24 Weighting is the process of adjusting a sample to infer results for the relevant population. To do this, a 'weight' is allocated to each sample unit - in this case, person (employee) records. The weight can be considered an indication of how many employees in the relevant population are represented by each person in the sample.

25 Estimates of the total number of persons with the specified characteristic should be obtained by summing the PERSON weights assigned to each linked record, using the variable called SWEIGHT.

26 Weights were calculated by calibrating to the following benchmarks:

- Total Earnings for each Statistical Area Level 4 by Occupation group (including 'Not known' SA4 and/or 'inadequately described' Occupation); and
- Total Earnings for each Age range by Sex by Occupation group (including 'inadequately described' Occupation).

27 This calibration ensures that the weighted sample estimates of total earnings in each of these groups match the total earnings for these groups according to the full Integrated Dataset.

Replicate Weights

28 Replicate weights can be used in the following manner to estimate the variance of the full sample statistic.

29 Using the replicate weights, sub-samples are repeatedly selected from the whole sample and the statistic of interest calculated for each of them. The variance of the full sample statistic is then estimated using the variability among the replicate statistics calculated from these sub-samples. The sub-samples are called 'replicate groups', and the statistics calculated from these replicate groups are called 'replicate estimates'.

30 The replicate weights for the Employee Earnings and Jobs microdata product were created using the jackknife method of replication. Each record in the Employee Earnings and Jobs microdata product has 60 replicate weights attached to it.

31 The formulae for calculating the SE and RSE of an estimate using the jackknife replicate weights are:

$$SE(y) = \sqrt{\frac{59}{60} \sum_{g=1}^{60} (y(g) - y)^2}$$

where g = 1,...,60 (the no. of replicate groups)

y(g) = weighted estimate, having applied the weights for replicate group g

y = weighted estimate from the full sample weight

$$RSE(y) \% = SE(y)/y * 100.$$

32 The 95% Margin of Error is calculated as MoE(y) = SE(y)*1.96.

33 This method can also be used when modelling relationships from unit record data. In modelling, the full sample would be used to estimate the parameter being studied (such as a regression coefficient) and the 60 replicate groups would be used to provide 60 replicate estimates of the survey parameter. The variance of the estimate of the parameter from the full sample is then approximated, as above, by the variability of the replicate estimates.

SOURCES OF ERROR

34 Potential sources of error, including sampling and non-sampling errors should be kept in mind when interpreting statistics from this product.

Sampling Error

35 Sampling error occurs because only a small proportion of the total population is used to produce estimates that represent the whole population. Sampling error refers to the fact that for a given sample size, each sample will produce different results, which will usually not be equal to the population value. Given the large sample size for the Employee Earnings and Jobs microdata product (1 in 10 employees), and stratified random sampling method used, sampling error will be relatively small in general, as quantified by the relative standard errors of estimates.

Non-sampling Error

36 Non-sampling error is caused by factors other than those related to using a sample in developing statistical outputs. It refers to the presence of any factor that would result in the data values not accurately reflecting the 'true' value for the population. They can occur at any stage of a collection (census, sample or administrative data) and are not easily identifiable or quantifiable.

37 The administrative data used in developing this microdata product is extensive in its scope, breadth, and utility, but it also contains missing and erroneous data, as well as data not suitable for the creation of official statistics without intervention. All these contribute to non-sampling errors. Simple editing strategies and cleaning have been applied to the administrative data used in this experimental output.

38 Non-sampling errors in this microdata product include but are not limited to those related to:

- **Linking accuracy**

During the construction of the Integrated Dataset a number of approaches were taken to allocate each ABN within a complex business structure to a single set of business characteristics. Further investigation into the allocation method is required as part of the future LEED development. For further detail on the accuracy of the linking, refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

- **Coverage**

Employees who did not report earnings on an ITR or did not receive an Individual PAYG summary from an employer were excluded from coverage in the Integrated Dataset. There were no businesses excluded on the basis of coverage.

- **Non-response**

This refers to blank fields in the PIT and PAYG forms received from the ATO. No attempt was made to impute the missing values as they were not sufficient to impact the analytical value of the dataset. For further details regarding missing values for key variables refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

- **Response errors**

This refers to a type of error caused by respondents intentionally or accidentally providing inaccurate responses. They are hard to detect and to quantify. The extent of occurrence of this error has not been assessed in this experimental exercise.

- **Processing errors**

This refers to errors that occur in the process of data collection, data entry, coding, editing and output. Once again, these are hard to identify and quantify and have not been assessed in this experimental exercise.

CONFIDENTIALITY

39 The **Census and Statistics Act, 1905** provides the authority for the ABS to collect statistical information, and requires that statistical output shall not be published or disseminated in a manner that is likely to enable the identification of a particular person or organisation. The confidentiality of respondents and businesses was maintained throughout the process. Access to taxation data is tightly controlled within the ABS. Policies and Guidelines governing the disclosure of information were implemented and followed in order to maintain the confidentiality of individuals and businesses.

40 Some techniques used to minimise the risk of identifying individuals and businesses in this microdata product are collapsing of categories (e.g., geography collapsed to state/territory level for the smaller states/territories of Tasmania, Northern Territory and Australian Capital Territory) and perturbation.

41 Perturbation involves making small random adjustment to values and is considered the most satisfactory technique for mitigating the risk of identification while maximising the range of information that can be released. The two earnings variables *Total earnings from all*

jobs held in reference period and *Gross payment amount per job held during the reference period* have been perturbed. Perturbation has had a negligible impact on the underlying distribution of the variables.

COHERENCE OF OUTPUTS ACROSS OTHER ABS COLLECTIONS

42 Analysis was conducted to assess the comparability of aggregate statistics produced from the full Integrated Dataset (experimental statistics) and those from related ABS household and business survey collections. They were found to be broadly coherent; however, differences were identified due to the differences in scope, sample design, collection methodology and processing approaches. Moreover, the Integrated Dataset is based on data collected for administrative purposes, whereas ABS collections are designed to create statistical outputs.

43 For further information on the coherence of the experimental statistics with ABS estimates refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

44 As the microdata product is a subset of the Integrated Dataset, similar differences between statistics produced from the microdata and those from other ABS surveys can be expected.

Glossary

GLOSSARY

Business File

Part of the Integrated Dataset. The Business File contains data from the Expanded Analytical Business Longitudinal Database extract for businesses which could be linked to a job on the Job File.

Client Dataset

This dataset contains detailed information about earnings, main occupation, tax withheld, deductions, and other items related to a single employee. This dataset is populated from information lodged through Individual Tax Returns to the Australian Taxation Office and feeds into the Employee File, a component of the Integrated Dataset.

Client Register

The register of client details maintained by the Australian Taxation Office, it is updated using information from the Individual Tax Returns. This file feeds into the Employee File, a component of the Integrated Dataset.

Employee

Persons who worked for a private or public sector employer and received pay for the reference period in the form of wages or salaries, a commission while also receiving a retainer, tips, piece rates or payments in kind. Persons who operated their own incorporated enterprises with or without hiring employees are also included as employees.

Employee File

Part of the Integrated Dataset. The Employee File contains data relating to each employee from the Personal Income Tax Client Register and Client Dataset.

Enterprise Group

A statistical unit covering all the operations in Australia of legal entities under common control. Multiple Australian Business Numbers can operate within a single Enterprise Group, and each Enterprise Group is broken up into one or more Type of Activity Units.

Individual Pay As You Go summary

The annual summary provided by an employer to the Australian Taxation Office with respect to an employee. It records job level information reported by employers about the gross payment made to an employee, tax withheld, and the start and end dates for each job. This also provides the Australian Business Number of the employer. This usually has a Tax File Number attached, although in some circumstances this may be missing or substituted for another code (e.g. if the employee did not provide it or is under the age of 18 and earns less than the tax-free threshold).

Individual Tax Return

The annual tax return submitted by individuals to the Australian Taxation Office.

Integrated Dataset

The physical file which constitutes the linked employer-employee data. The Integrated Dataset is comprised of three main subsets:

- Employee File;
- Job File; and
- Business File.

Job

It is defined as a link between an employee and a business for \$1 or more in payment as reported on an Individual Pay As You Go summary. An employee can have multiple jobs with the same or different businesses during the financial year, and can hold two or more jobs concurrently.

Job File

Part of the Integrated Dataset. The Job File contains data relating to each job from the Personal Income Tax Individual Pay As You Go dataset.

Non-profiled population (simple businesses)

The majority of businesses have simple structures and the unit registered for an ABN will satisfy ABS statistical reporting requirements. These businesses form the non-profiled population.

Profiled population (complex businesses)

For those businesses where the ABN is not considered suitable for ABS statistical requirements, the ABS maintains its own units structure through direct contact with the business. This population consists typically of large, diverse and complex structured businesses, and constitute the profiled population.

For further information on Non-profiled and Profiled population of businesses refer to the *Information Paper: Construction of Experimental Statistics on Employee Earnings and Jobs from Administrative Data, Australia, 2011-12* (cat. no. 6311.0).

Statistical Area Level 4

An area defined in the Australian Statistical Geography Standard and designed for the output of labour force data and to reflect labour markets. In rural areas, SA4s generally represent aggregations of multiple small labour markets with socioeconomic connections or

similar industry characteristics. Large regional city labour markets are generally defined by a single SA4. Within major metropolitan labour markets SA4s represent sub-labour markets. SA4s generally have a population over 100,000 people to enable accurate labour force survey data to be generated. There are 88 SA4s and they cover the whole of Australia without gaps or overlaps.

For further information, refer to Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas (cat. no. 1270.0.55.001).

Start and End Dates

Start and end dates associated with each job as reported on an Individual Pay As You Go summary.

Type of Activity Unit

The TAU is a producing unit comprising one or more business entities, sub-entities or branches of a business entity that can report production and employment activities via a minimum set of data items. The activity of the unit should homogenous as far as possible.

Abbreviations

ABBREVIATIONS

ABN	Australian Business Number
ABS	Australian Bureau of Statistics
ANZSIC	Australian and New Zealand Standard Industrial Classification
ATO	Australian Taxation Office
EABLD	Expanded Analytical Business Longitudinal Database
EEJ	Employee Earnings and Jobs
EG	Enterprise Group
ITR	Individual Tax Return
MoE	Margin of Error
PAYG	Pay As You Go
PIT	Personal Income Tax
RSE	Relative Standard Error
SA4	Statistical Area Level 4
SE	Standard Error
TAU	Type of Activity Unit
TFN	Tax File Number

Data Cubes (I-Note) - Data Cubes

A test file has been created for the Employee Earnings and Jobs microdata product. The

purpose of the test file is to allow researchers/analysts to become familiar with the data structure and prepare code/programs prior to applying for, or commencing, an ABS Data Laboratory session.

The test file is available as a free download from the Downloads tab. It can also be made available in other file formats on request, if required. For further information users should email microdata.access@abs.gov.au or telephone (02) 6252 7714.

The test file does not contain real data, and cannot be used for analysis.

© Commonwealth of Australia

All data and other material produced by the Australian Bureau of Statistics (ABS) constitutes Commonwealth copyright administered by the ABS. The ABS reserves the right to set out the terms and conditions for the use of such material. Unless otherwise noted, all material on this website – except the ABS logo, the Commonwealth Coat of Arms, and any material protected by a trade mark – is licensed under a Creative Commons Attribution 2.5 Australia licence